AD-A061 337 CORNELL UNIV ITHACA N Y SCHOOL OF OPERATIONS RESEARC--ETC F/G 12/1
A TWO-STAGE MINIMAX PROCEDURE WITH SCREENING FOR SELECTING THE --ETC(U)
OCT 78 A C TAMHANE, R E BECHHOFER DAAG29-77-C-0003

UNCLASSIFIED TR-377 NL

1 OF 1
AD
A061337

END
DATE
FILMED
2 - 79
DDC

LEVEL II

(6)

A057823

SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING

COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK

(9) TECHNICAL REPORT NO. 377

June 1978

(11) (Revised October 1978)

(6) A TWO-STAGE MINIMAX PROCEDURE WITH SCREENING FOR SELECTING THE LARGEST NORMAL MEAN (II) AN IMPROVED PCS LOWER BOUND AND ASSOCIATED TABLES.

by

(10) Ajit C. Tamhane
Northwestern University

Robert E. Bechhofer
Cornell University

(12) 33 p.

(14) TR-377

Prepared under
NSF Grant ENG 77-06112
at Northwestern University
and
(15) U.S. Army Research Office-Durham contract DAAG29-77-C-0003,
Office of Naval Research contract N00014-75-C-0586
at Cornell University

THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS

TABLE OF CONTENTS

Abstract

1. Introduction and summary 1

2. Preliminaries 3

   2.1 Assumptions 3

   2.2 Goal and probability requirement 3

   2.3 Two-stage procedure (P_2) 3

   2.4 U-minimax design criterion 5

3. An improved lower bound on $P_{\underline{\mu}}\{CS|P_2\}$, and the associated U-minimax optimization problem 5

   3.1 Improved lower bound on $P_{\underline{\mu}}\{CS|P_2\}$ 5

   3.2 U-minimax optimization problem 8

4. Constants to implement $P_2(C_1)$ for $k \geq 3$ 10

   4.1 Tables of constants to implement $P_2(C_1)$ 10

   4.2 Details of computations of constants 10

5. The performance of $P_2(\cdot)$ relative to P_1 15

   5.1 $P_2(E)$, $P_2(C_1)$ and $P_2(C_3)$ vs. P_1 for $k = 2$ 18

   5.2 $P_2(C_1)$ vs. P_1 for $k \geq 3$ 18

   5.3 $P_2(C_1)$, $P_2(C_2)$ and $P_2(C_3)$ vs. P_1 for $k \geq 3$ 20

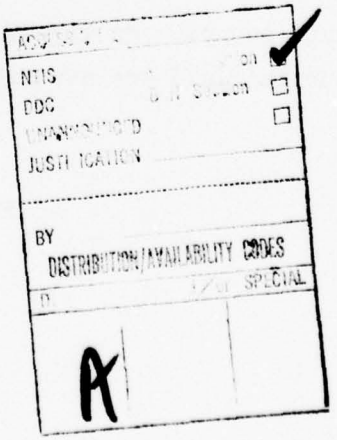
6. Numerical example 22

7. Concluding remarks 24

8. Directions of future research 24

9. Acknowledgments 24

   Bibliography 25



A TWO-STAGE MINIMAX PROCEDURE WITH SCREENING FOR SELECTING THE LARGEST NORMAL MEAN (II): AN IMPROVED PCS LOWER BOUND AND ASSOCIATED TABLES

Ajit C. Tamhane

Northwestern University, Evanston, Illinois

Robert E. Bechhofer

Cornell University, Ithaca, New York

ABSTRACT

This paper is a follow-up to an earlier article by the authors in which they proposed a two-stage procedure with screening to select the normal population with the largest population mean when the populations have a common known variance. The two-stage procedure has the highly desirable property that the expected total number of observations required by the procedure is always less than the total number of observations required by the corresponding single-stage procedure of Bechhofer (1954), regardless of the configuration of the population means. The present paper contains new results which make possible the more efficient implementation of the two-stage procedure. Tables for this purpose are given, and the improvements achieved (which are substantial) are assessed.

1. INTRODUCTION AND SUMMARY

The present paper is a follow-up to Tamhane and Bechhofer (1977) (henceforth referred to as T-B) and contains some new results which make possible the more efficient implementation of the two-stage procedure proposed in Section 4 of T-B. In order to make the present paper somewhat self contained, certain results from T-B are repeated here (without proof); the reader is referred to T-B for background and motivation as well as for the necessary proofs.

In T-B we studied in depth a two-stage procedure (P_2) for selecting the largest normal mean. This procedure (which employs the indifference-zone approach of Bechhofer (1954)) screens out "noncontending" populations in the first stage and selects the "best" population from among the "contending" populations which enter the second stage. In order to determine the constants necessary to implement P_2, we proposed in T-B the criterion of minimizing the maximum (over the entire parameter space) of the expected total sample size required by P_2 subject to the procedure's guaranteeing a specified probability requirement. As a consequence, P_2 based on this unrestricted minimax (U-minimax) design criterion possesses the highly desirable property that the expected total sample size required by P_2 is always less than the total sample size required by the best competing single-stage procedure (P_1) of Bechhofer (1954), regardless of the true configuration of the population means.

As noted in Section 10 of T-B, there were two main unsolved problems associated with P_2 applied to three or more $(k \geq 3)$ populations. First, the so-called least-favorable configuration (LFC) of the population means has not yet been determined for $k \geq 3$; knowledge of this configuration is required in order to determine the best set of constants necessary to implement P_2. Second, even if the LFC of the population means were known for $k \geq 3$, the problem of evaluating the probability of a correct

selection (P{CS}) associated with P_2 (see (5.1) of T-B) when the population means are in that configuration would still remain; it is extremely difficult and costly to evaluate the exact P{CS} associated with P_2 on a computer, even if the population means are in the so-called "slippage" configuration (see (5.6) of T-B).

It is possible to determine a set of constants (although not the best set) to implement P_2 if a lower bound to the P{CS} of P_2 can be found and the LFC of the population means determined for that lower bound; such a set of constants provides a conservative solution to the problem. It was this device that we adopted in T-B in order to circumvent the first unsolved problem; it turned out that it was straightforward to determine the LFC of the population means for the lower bound that we adopted (see (5.8) of T-B), and in addition the integrals associated with that lower bound proved to be very easy to compute.

A referee of T-B proposed a new lower bound to the P{CS} of P_2 (see Section 11 of T-B), his bound being uniformly superior to ours; it is also straightforward to determine the LFC of the population means relative to the referee's lower bound, and the resulting integrals are easy to evaluate. Using this new lower bound we could have computed a new set of constants to implement P_2, and thereby obtained a less conservative solution to our problem.

In the present paper we obtain a third lower bound to the P{CS} of P_2 --one which is uniformly superior to the referee's, and relative to which the LFC of the population means is easily obtained. However, it is quite a bit more difficult and costly (although not prohibitively so) to evaluate the resulting function than was the case for our original bound or for the referee's. It turned out that such computations were justified since results obtained with this newest bound yield a significant improvement over all of our previous results. We make these ideas precise in the next sections. The comparisons between our new results and our previous ones are made in Section 5. In Section 6 a

numerical example is given which illustrates the options (in terms of choice among procedures) available to the experimenter, and the strikingly different consequences associated with each option.

2. PRELIMINARIES

2.1 Assumptions

Let Π_i $(1 \leq i \leq k)$ denote a normal population with unknown mean μ_i and known variance σ^2, and let Ω denote the parameter space of vectors $\underline{\mu} = (\mu_1,\ldots,\mu_k)$. Denote the ranked values of the μ_i by $\mu_{[1]} \leq \cdots \leq \mu_{[k]}$, and let $\delta_{i,j} = \mu_{[i]} - \mu_{[j]}$. We assume no prior knowledge concerning the pairing of the Π_i with the $\mu_{[j]}$ $(1 \leq i,j \leq k)$. Any one of the populations (if there is more than one) with μ-value equal to $\mu_{[k]}$ is regarded as "best."

2.2 Goal and Probability Requirement

The goal of the experimenter is to select a best population. This event is referred to as a correct selection (CS). The experimenter restricts consideration to procedures (P) which guarantee the probability requirement

$$P_{\underline{\mu}}\{CS|P\} \geq P^* \quad \text{for all} \quad \underline{\mu} \in \Omega(\delta^*) \tag{2.1}$$

where $\{\delta^*,P^*\}$ $0 < \delta^* < \infty$, $1/k < P^* < 1$ are specified prior to experimentation, and $\Omega(\delta^*) = \{\underline{\mu} \in \Omega | \delta_{k,k-1} \geq \delta^*\}$.

2.3 Two-stage Procedure (P_2)

In T-B we proposed a two-stage procedure $P_2 = P_2(n_1,n_2,h)$ (previously considered by Alam (1970)) which depends on nonnegative integers n_1,n_2 and a real constant $h \geq 0$. The constants (n_1,n_2,h) depend on k and $\{\delta^*,P^*\}$, and are chosen so that P_2 guarantees (2.1) and possesses a certain minimax property (given by 2.2)).

Procedure P_2

Stage 1: Take a random sample of size n_1 from each Π_i and

compute the sample mean $\overline{X}_i^{(1)}$ $(1 \leq i \leq k)$. Let $\overline{X}_{[k]}^{(1)} = \max_{1\leq i\leq k} \overline{X}_i^{(1)}$. Determine the subset I of $\{1,2,\ldots,k\}$ where $I = \{i \mid \overline{X}_i^{(1)} \geq \overline{X}_{[k]}^{(1)} - h\}$. The populations $\Pi_1,\ldots,\Pi_k$ with subscripts in I are the ones which enter the second stage (if I has more than one element). Let Π_I denote the set of populations with subscripts in I.

a) If Π_I contains only one population, stop sampling and assert that the population associated with $\overline{X}_{[k]}^{(1)}$ is best.

b) If Π_I contains more than one population, proceed to the second stage.

Stage 2: Take a random sample of size n_2 from each Π_i with $i \in I$ and compute the sample mean $\overline{X}_i^{(2)}$. Compute the cumulative sample mean $\overline{X}_i = (n_1\overline{X}_i^{(1)} + n_2\overline{X}_i^{(2)})/(n_1+n_2)$ for each Π_i with $i \in I$. Assert that the population associated with $\overline{X}_{[k]} = \max_{i\in I} \overline{X}_i$ is best.

Remark 2.1: A two-stage procedure of Somerville (1974) which is related to ours has recently come to our attention. His procedure eliminates a predetermined number of populations at the end of the first stage whereas ours eliminates a random number; thus for favorable configurations of the population means, his procedure always requires two stages and a fixed total number of observations, whereas ours often requires only a single stage or two stages with a small total number of observations. Moreover, after the first-stage data are used to determine which populations enter the second stage, Somerville's procedure ignores the information concerning the population means obtained in the first stage. Our procedure not only uses the first stage data to determine which populations enter the second stage, but also for those populations which do enter the second stage our procedure pools the first stage and second stage information concerning the associated population means; thus our procedure makes fuller use of the information in the total experiment.

Let T denote the total sample size and $E_{\underline{\mu}}\{T|P_2\}$ denote the expected total sample size required by P_2. In T-B (Section 4.2) we proposed the following unrestricted minimax (U-minimax) design criterion to determine (n_1,n_2,h) guaranteeing (2.1).

2.4 U-minimax Design Criterion

For given k and specified $\{\delta^*,P^*\}$ choose (n_1,n_2,h) to

$$\text{minimize} \quad \sup_{\underline{\mu}\in\Omega} E_{\underline{\mu}}\{T|P_2\}$$
$$\text{subject to} \quad \inf_{\underline{\mu}\in\Omega(\delta^*)} P_{\underline{\mu}}\{CS|P_2\} \geq P^*, \tag{2.2}$$

where n_1,n_2 are nonnegative integers and $h \geq 0$. We denote by $(n_1,n_2,h|E)$ the exact solution to (2.2), and by $P_2(E)$ the procedure using that solution. If a lower bound on $P_{\underline{\mu}}\{CS|P_2\}$ is used in the l.h.s. of the constraint of (2.2), then we denote by $(n_1,n_2,h|C)$ the corresponding conservative solution to (2.2), and by $P_2(C)$ the procedure using that solution.

3. AN IMPROVED LOWER BOUND ON $P_{\underline{\mu}}\{CS|P_2\}$, AND THE ASSOCIATED U-MINIMAX OPTIMIZATION PROBLEM

3.1 Improved Lower Bound on $P_{\underline{\mu}}\{CS|P_2\}$

Our new lower bound is given in the following theorem which is proved below.

Theorem 3.1: For all $\underline{\mu} \in \Omega$ we have the following inequalities:

$$P_{\underline{\mu}}\{CS|P_2\} > A > B > C$$

where

$$A = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi_2\left[\frac{(\delta_{k,i}+h)\sqrt{n_1}}{\sigma} - x_1, \frac{\delta_{k,i}\sqrt{m}}{\sigma} - x_2 \middle| \sqrt{p}\right]$$
$$\times\, d\Phi_2[x_1,x_2|\sqrt{p}], \tag{3.1a}$$

$$B = \left\{\int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[x + (\delta_{k,i}+h)\sqrt{n_1}/\sigma]d\Phi(x)\right\}$$

$$\times\left\{\int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[x + \delta_{k,i}\sqrt{m}/\sigma]d\Phi(x)\right\}, \qquad (3.1b)$$

$$C = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[x + (\delta_{k,i}+h)\sqrt{n_1}/\sigma]d\Phi(x)$$

$$+ \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[x + \delta_{k,i}\sqrt{m}/\sigma]d\Phi(x) - 1. \qquad (3.1c)$$

Here $\Phi_2(\cdot,\cdot|\rho)$ denotes the standard bivariate normal cdf with correlation coefficient ρ $(-1 < \rho < 1)$; $\Phi(\cdot)$ denotes the standard univariate normal cdf, $m = n_1 + n_2$, and $p = n_1/m$.

Proof: Let $\bar{X}_{(i)}^{(1)}$ and $\bar{X}_{(i)}$ denote the first stage sample mean and the cumulative (first stage plus second stage) sample mean, respectively, from the population with mean $\mu_{[i]}$ $(1 \leq i \leq k)$. Then

$$P_{\underline{\mu}}\{CS|P_2\} \geq P_{\underline{\mu}}\{\bar{X}_{(k)}^{(1)} \geq \bar{X}_{(i)}^{(1)} - h,\ \bar{X}_{(k)} > \bar{X}_{(i)}\ (1 \leq i \leq k-1)\} \qquad (3.2)$$

$$= P\left\{U_i \leq \frac{(\delta_{k,i}+h)}{\sigma}\sqrt{\frac{n_1}{2}},\ V_i \leq \frac{\delta_{k,i}}{\sigma}\sqrt{\frac{m}{2}}\ (1 \leq i \leq k-1)\right\}$$

where $U_i = (\bar{X}_{(i)}^{(1)} - \bar{X}_{(k)}^{(1)} + \delta_{k,i})\sqrt{n_1}/\sigma\sqrt{2}$ and $V_i = (\bar{X}_{(i)} - \bar{X}_{(k)} + \delta_{k,i})\sqrt{m}/\sigma\sqrt{2}$ $(1 \leq i \leq k-1)$. We note that the U_i and V_i are standard normal r.v.'s with $\mathrm{Corr}\{U_i,U_j\} = \mathrm{Corr}\{V_i,V_j\} = 1/2$, $\mathrm{Corr}\{U_i,V_i\} = \sqrt{p}$, and $\mathrm{Corr}\{U_i,V_j\} = \sqrt{p}/2$ $(i \neq j;\ 1 \leq i,j \leq k-1)$; thus using the representation (see Bechhofer and Tamhane (1974)) $(U_i,V_i)' = (Y_{1i}+Y_{10}, Y_{2i}+Y_{20})'/\sqrt{2}$ $(1 \leq i \leq k-1)$ where the $(Y_{1j},Y_{2j})'$ $(0 \leq j \leq k-1)$ are i.i.d. $N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \sqrt{p}\\ \sqrt{p} & 1\end{pmatrix}\right)$ we obtain A of (3.1a) which is our new lower bound. The referee's lower bound B is obtained by replacing all of the

correlations between the U_i and the V_j $(1 \leq i,j \leq k-1)$ by zeros and applying Slepian's inequality. The inequality between B and our original lower bound C was shown in Section 11 of T-B.

We note that for A, B, and C the LFC of the population means, i.e., the configuration of the population means which minimizes A, B, and C subject to $\underline{\mu} \in \Omega(\delta^*)$, is given by $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ (which is also the conjectured LFC for the exact $P_{\underline{\mu}}\{CS|R_2\}$ subject to $\underline{\mu} \in \Omega(\delta^*)$). Thus we now obtain

Corollary: For all $\underline{\mu} \in \Omega(\delta^*)$ we have the following inequalities:

$$P_{\underline{\mu}}\{CS|R_2\} > A(\delta^*) > B(\delta^*) > C(\delta^*) \tag{3.3}$$

where

$$A(\delta^*) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi_2^{k-1}\left[\frac{(\delta^*+h)\sqrt{n_1}}{\sigma} - x_1, \frac{\delta^*\sqrt{m}}{\sigma} - x_2 \middle| \sqrt{\rho}\right] \times d\Phi_2[x_1,x_2|\sqrt{\rho}], \tag{3.4a}$$

$$B(\delta^*) = \left\{\int_{-\infty}^{\infty} \Phi^{k-1}[x + (\delta^*+h)\sqrt{n_1}/\sigma]d\Phi(x)\right\} \times \left\{\int_{-\infty}^{\infty} \Phi^{k-1}[x + \delta^*\sqrt{m}/\sigma]d\Phi(x)\right\}, \tag{3.4b}$$

$$C(\delta^*) = \int_{-\infty}^{\infty} \Phi^{k-1}[x + (\delta^*+h)\sqrt{n_1}/\sigma]d\Phi(x) + \int_{-\infty}^{\infty} \Phi^{k-1}[x + \delta^*\sqrt{m}/\sigma]d\Phi(x) - 1. \tag{3.4c}$$

In the optimization problem (2.2) we shall replace the constraint by $A(\delta^*) \geq P^*$; we denote by $(n_1,n_2,h|C_1)$ the corresponding conservative solution and by $R_2(C_1)$ the associated

procedure. Similarly, conservative solutions and procedures result if we employ the constraints $B(\delta^*) \geq P^*$ or $C(\delta^*) \geq P^*$ in (2.2) obtaining $(n_1,n_2,h|C_2)$ and $P_2(C_2)$, and $(n_1,n_2,h|C_3)$ and $P_2(C_3)$, respectively.

Remark 3.1: If we let $h \to \infty$ in A, B, or C of (3.1a), (3.1b), (3.1c), respectively, we obtain in each case $\int_{-\infty}^{\infty} \Pi_{i=1}^{k-1} \Phi[x + \delta_{k,i}\sqrt{m}/\sigma]d\Phi(x)$ which is an expression for $P_{\underline{\mu}}\{CS|P_1\}$ where P_1 uses a common single-stage sample size m per population. Therefore P_1 is a special case of any P_2 based on either (3.1a) or (3.1b) or (3.1c). Now it was shown in Bechhofer (1954) that

$$\inf_{\underline{\mu}\in\Omega(\delta^*)} P_{\underline{\mu}}\{CS|P_1\} = \int_{-\infty}^{\infty} \Phi^{k-1}[x + \delta^*\sqrt{m}/\sigma]d\Phi(x). \qquad (3.5)$$

Thus if we let n denote the smallest value of m for which the r.h.s. of (3.5) is equal to or greater than P^* (i.e., n is the smallest single-stage sample size that guarantees (2.1)), then for $i = 1,2,3$ we have

$$E_{\underline{\mu}}\{T|P_2(C_i)\} \leq kn \qquad (3.6)$$

for all $\underline{\mu} \in \Omega$.

3.2 U-minimax optimization problem

Before we state our optimization problem we cite the following results given by Theorems 6.1 and 6.2, respectively, of T-B:
(1) For any $\underline{\mu} \in \Omega$ we have

$$E_{\underline{\mu}}\{T|P_2\} = kn_1 + n_2 \sum_{i=1}^{k} \int_{-\infty}^{\infty} \left\{ \prod_{\substack{j=1 \\ j\neq i}}^{k} \Phi\left[x + \frac{(\delta_{i,j}+h)\sqrt{n_1}}{\sigma}\right] - \prod_{\substack{j=1 \\ j\neq i}}^{k} \Phi\left[x + \frac{(\delta_{i,j}-h)\sqrt{n_1}}{\sigma}\right] \right\} d\Phi(x), \qquad (3.7)$$

and

$$(2)\quad \sup_{\underline{\mu}\in\Omega} E_{\underline{\mu}}\{T|P_2\} = kn_1 + kn_2 \int_{-\infty}^{\infty} \{\Phi^{k-1}[x + h\sqrt{n_1}/\sigma] - \Phi^{k-1}[x - h\sqrt{n_1}/\sigma]\}d\Phi(x) \tag{3.8}$$

which occurs when $\mu_{[1]} = \mu_{[k]}$ (referred to as the equal means configuration (EMC)).

It is more convenient to work with continuous variables than with discrete variables. Thus we define new design constants

$$c_1 = \frac{\delta^*\sqrt{n_1}}{\sigma},\ c_2 = \frac{\delta^*\sqrt{n_2}}{\sigma},\ d = \frac{h\sqrt{n_1}}{\sigma} \tag{3.9}$$

which we regard as nonnegative <u>continuous</u> variables. Then the design constants $(n_1,n_2,h|C_1)$ can be approximated by solving the following <u>continuous</u> optimization problem:

For given k and specified P^* choose (c_1,c_2,d) to

$$\text{minimize } kc_1^2 + kc_2^2 \int_{-\infty}^{\infty} \{\Phi^{k-1}(x+d) - \Phi^{k-1}(x-d)\}d\Phi(x)$$

$$\text{subject to } \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi_2^{k-1}[c_1+d-x_1,\ \sqrt{c_1^2+c_2^2}-x_2|\sqrt{p}] \times d\Phi_2(x_1,x_2|\sqrt{p}) \geq P^* \tag{3.10}$$

where $p = c_1^2/(c_1^2+c_2^2)$ and $c_1,c_2,d \geq 0$. We denote the solution to (3.10) by $(\hat{c}_1,\hat{c}_2,\hat{d}|C_1)$ and for specified δ^* use the <u>approximate</u> design constants

$$\hat{n}_1 = \left[\left(\frac{\hat{c}_1\sigma}{\delta^*}\right)^2\right],\ \hat{n}_2 = \left[\left(\frac{\hat{c}_2\sigma}{\delta^*}\right)^2\right],\ \hat{h} = \frac{\hat{d}\delta^*}{\hat{c}_1}, \tag{3.11}$$

where $[z]$ denotes the smallest integer $\geq z$, to implement $P_2(C_1)$.

For $k = 2$ we note that (3.10) can be written as:

For specified P^* choose (c_1,c_2,d) to

$$\text{minimize } 2c_1^2 + 2c_2^2\{\Phi(d/\sqrt{2}) - \Phi(-d/\sqrt{2})\}$$

$$\text{subject to } \Phi_2[(c_1+d)/\sqrt{2},\ \sqrt{c_1^2+c_2^2}/\sqrt{2}|\sqrt{p}] \geq P^*. \tag{3.12}$$

4. CONSTANTS TO IMPLEMENT $P_2(C_1)$ FOR $k \geq 3$

4.1 Tables of Constants to Implement $P_2(C_1)$

Table I contains constants $(\hat{c}_1,\hat{c}_2,\hat{d}|C_1)$ necessary to approximate $(n_1,n_2,h|C_1)$ for $k = 2$ and selected P^*; these constants are the solutions to (3.12). Also included in Table I are the constants $(\hat{c}_1,\hat{c}_2,\hat{d}|E)$ and $(\hat{c}_1,\hat{c}_2,\hat{d}|C_3)$ necessary to implement the exact procedure and $P_2(C_3)$, respectively; these were given in T-B and Tamhane (1975), respectively. We include the latter constants here in order for the reader to see how the constants $(\hat{c}_1,\hat{c}_2,\hat{d}|C_1)$ compare with them; however, we emphasize that in practice one would only use $(\hat{c}_1,\hat{c}_2,\hat{d}|E)$ since these constants are optimal.

Table II contains constants $(\hat{c}_1,\hat{c}_2,\hat{d}|C_1)$ necessary to approximate $(n_1,n_2,h|C_1)$ for $k = 3(1)10,12,15,25$ and selected P^*; these constants are the solutions to (3.10). For fixed (large) P^* we have found that $\hat{c}_1^2$ and $\hat{c}_2^2$ are approximately linear in $\log_e(k-1)$; similarly, for fixed k and (large) P^*, $\hat{c}_1^2$ and $\hat{c}_2^2$ are approximately linear in $\log_e\{P^*/(1-P^*)\}$. We have not been able to characterize the behavior of $\hat{d}$ in a simple way. The constants given in Table II for $k = 6(1)9$ and 12 were obtained by quadratic interpolation of $\hat{c}_1^2$, $\hat{c}_2^2$ and $\hat{d}^2$ against $\log_e(k-1)$ for fixed $P^* = 0.75$, 0.90, 0.95 and 0.99; these constants (particularly $\hat{d}$) are not as accurate as the ones which served as the basis for the interpolation.

4.2 Details of Computations of Constants

All of the computations were carried out on Northwestern's CDC 6600 computer in 32-bit arithmetic. The generalized reduced gradient (GRG) algorithm of Abadie and Carpentier (1969) was used to solve the constrained nonlinear optimization problems (3.10) and (3.12). The constants $(\hat{c}_1,\hat{c}_2,\hat{d}|C_3)$ given in T-B were used as initial guesses in the GRG algorithm; even with these relatively "good" guesses, at least 10 and often more iterations were

TABLE I

Constants to Implement $P_2(E)$, $P_2(C_1)$ and $P_2(C_3)$ for k = 2

P*	$(\hat{c}_1,\hat{c}_2,\hat{d}\|E)$			$(\hat{c}_1,\hat{c}_2,\hat{d}\|C_1)$			$(\hat{c}_1,\hat{c}_2,\hat{d}\|C_3)$		
	$\hat{c}_1$	$\hat{c}_2$	$\hat{d}$	$\hat{c}_1$	$\hat{c}_2$	$\hat{d}$	$\hat{c}_1$	$\hat{c}_2$	$\hat{d}$
0.9999	4.540	2.909	0.9722	4.508	2.887	1.118	4.480	3.059	1.120
0.9995	3.974	2.671	0.9562	3.958	2.675	1.043	3.909	2.836	1.135
0.999	3.706	2.571	0.9482	3.698	2.517	1.083	3.640	2.727	1.153
0.99	2.719	2.091	0.9191	2.708	2.049	1.097	2.623	2.254	1.301
0.95	1.862	1.616	0.8807	1.869	1.525	1.161	1.739	1.665	1.842
0.90	1.427	1.313	0.8528	1.454	1.204	1.173	1.308	1.288	2.599
0.85	1.139	1.093	0.8417	1.193	0.9159	1.311			
0.80	0.9158	0.9097	0.8270	0.9627	0.7568	1.360			
0.75	0.7280	0.7416	0.8200	0.7987	0.5595	1.363			
0.70	0.5624	0.5866	0.8078	0.6290	0.4155	1.451			
0.65	0.4123	0.4330	0.8020	0.4522	0.3132	1.759			
0.60	0.2698	0.2878	0.7971	0.2998	0.2003	1.916			
0.55	0.1337	0.1432	0.7914	0.1557	0.08813	2.100			

TABLE II

Constants to Implement $P_2(C_1)$ for $k \geq 3$

k	P*	$(\hat{c}_1, \hat{c}_2, \hat{d} \mid C_1)$		
		$\hat{c}_1$	$\hat{c}_2$	$\hat{d}$
3	0.99	2.791	2.406	1.205
	0.95	1.999	1.846	1.552
	0.90	1.578	1.525	2.100
	0.75	0.9986	0.9485	3.989
4	0.99	2.965	2.508	1.222
	0.95	2.139	2.090	1.452
	0.90	1.760	1.777	1.700
	0.75	1.161	1.128	3.545
5	0.99	3.043	2.693	1.241
	0.95	2.252	2.257	1.362
	0.90	1.845	1.963	1.537
	0.75	1.260	1.277	2.821
6	0.99	3.087	2.815	1.261
	0.95	2.307	2.397	1.352
	0.90	1.916	2.121	1.478
	0.75	1.332	1.440	2.556
7	0.99	3.130	2.909	1.278
	0.95	2.355	2.505	1.332
	0.90	1.969	2.240	1.398
	0.75	1.386	1.582	2.276
8	0.99	3.163	2.989	1.294
	0.95	2.394	2.595	1.321
	0.90	2.011	2.340	1.342
	0.75	1.428	1.709	2.049
9	0.99	3.189	3.060	1.309
	0.95	2.425	2.673	1.315
	0.90	2.045	2.426	1.304
	0.75	1.463	1.824	1.861
10	0.99	3.194	3.142	1.322
	0.95	2.452	2.744	1.322
	0.90	2.067	2.507	1.342
	0.75	1.500	1.889	1.570

TABLE II (continued)

12	0.99	3.243	3.231	1.349
	0.95	2.492	2.858	1.318
	0.90	2.120	2.630	1.256
	0.75	1.541	2.100	1.468
15	0.99	3.272	3.376	1.384
	0.95	2.532	3.001	1.346
	0.90	2.174	2.791	1.330
	0.75	1.588	2.350	1.364
	0.60	1.235	1.667	1.979
25	0.99	3.340	3.649	1.463
	0.95	2.621	3.302	1.411
	0.90	2.271	3.121	1.358
	0.75	1.704	2.809	1.256
	0.60	1.360	2.732	1.099
	0.45	1.014	2.227	1.219

required to arrive in the neighborhood of the absolute optimum. The objective function is relatively flat in the region of the absolute optimum which results in extremely slow convergence in the later iterations. A maximum limit of 25 was placed on the number of iterations, and the algorithm was terminated sooner if no change was observed in the first four significant digits of the objective function in five successive iterations. Thus we would expect our solutions to be reasonably close to the absolute optima.

Each iteration of the GRG algorithm corresponds to at least one (and often many) evaluations of the double integral in the constraint of (3.10) and its partial derivatives with respect to c_1, c_2 and d. We evaluated these partials numerically by taking the value of the double integral at the current solution (c_1,c_2,d) as the base value, say $\psi(c_1,c_2,d)$, and approximated the derivatives as: $\partial\psi/\partial c_1 \simeq \{\psi(c_1+\Delta,c_2,d) - \psi(c_1,c_2,d)\}/\Delta$, etc., where $\Delta = 10^{-4}$. Since it was necessary to evaluate the double integral in the constraint of (3.10) at least four and

often many more times in each iteration of the GRG algorithm, a fast and accurate method was sought for this purpose.

We used Monte Carlo (MC) sampling to estimate the value of this double integral by noting that it equals

$$E\{\Phi_2^{k-1}[c_1+d-X_1,\ \sqrt{c_1^2+c_2^2}-X_2 \mid \sqrt{p}]\} \qquad (4.1)$$

where X_1 and X_2 are standard normal r.v.'s with $\mathrm{Corr}\{X_1,X_2\} = \sqrt{p} = c_1/\sqrt{c_1^2+c_2^2}$. Each estimate of (4.1) was based on an average over 200 runs. In each run (X_1,X_2) was generated by first generating a pair (Z_1,Z_2) of independent standard normal r.v.'s using the Box-Müller algorithm and then employing the transformation $X_1 = Z_1$, $X_2 = \sqrt{p}\, Z_1 + \sqrt{1-p}\, Z_2$. The Fortran library program RANF was used to generate the uniform [0,1] r.v.'s needed as inputs for the Box-Müller algorithm.

We evaluated Φ_2 using Borth's (1973) modification of Owen's (1958) method. This modification is based on the fact that subject to a specified accuracy, Owen's method is fast only for a certain range of values of the parameters of Φ_2; for other values of the parameters a computing method proposed by Borth is faster. The modified method which is a composite of the two methods is thus faster than Owen's method. For the Owen method part we used the IMSL subroutine MDBNOR for which a limit on the maximum error in evaluating Φ_2 is specified to be 10^{-5} in the IMSL manual. For the Borth method part the limit on the maximum error in evaluating a certain T-function necessary to obtain Φ_2 (see equation (2) of Borth (1973)) is specified to be 10^{-7} in his article. In addition to the T-function it is also necessary to compute some standard normal cdf values to obtain Φ_2. (All of these steps are carried out internally in MDBNOR.) We used the approximation to Φ given by equation 26.2.17 of Abramowitz and Stegun (1964) which is accurate to within $\pm 7.5 \times 10^{-8}$. Thus the overall accuracy in the evaluation of Φ_2 may be estimated to be $\pm 10^{-5}$. (We mention that we tried to use Cadwell's (1951)

approximation to Φ_2 but found its accuracy to be unacceptable for our purposes.)

For evaluating the double integral we also tried using the Gauss-Legendre quadrature method (with the integrals over $-\infty$ to $+\infty$ approximated by integrals over -6 to +6) with 16 nodes, i.e., 256 integrand evaluations. Although this method requires slightly more computer time than the MC method with 200 runs, we found that in general the MC method gave a more accurate estimate of the double integral; hence we adopted the MC method.

For evaluating the single integral in the objective function of (3.10) we used the Romberg quadrature method for which the maximum error was controlled at 10^{-5}. Φ appearing in the integrand was evaluated using the approximation given in Abramowitz and Stegun (1964) referred to earlier.

The reported values of $\hat{c}_1$, $\hat{c}_2$, $\hat{d}$ are rounded off in the fourth significant digit and are estimated to be correct to at least the first three significant digits.

5. THE PERFORMANCE OF $P_2(\cdot)$ RELATIVE TO P_1

As a measure of the efficiency of P_1 (Bechhofer (1954)) relative to that of P_2 when both guarantee the same probability requirement (2.1), we consider the ratio (termed relative efficiency (RE(P_1:P_2))

$$E_{\underline{\mu}}\{T|P_2\}/kn, \tag{5.1}$$

where $n = [(\hat{c}\sigma/\delta^*)^2]$, and $\hat{c}$ is the solution of

$$\int_{-\infty}^{\infty} \Phi^{k-1}(x+c)d\Phi(x) = P^*. \tag{5.2}$$

Clearly RE depends on $\underline{\mu}$ and $\{\delta^*,P^*\}$; values of RE less than one favor P_2 over P_1. To remove the dependence on δ^* we use the continuous approximations to $E_{\underline{\mu}}\{T|P_2\}$ and n (thereby ignoring the fact that the sample sizes must be integers). RE is then given by

$$\left\{ kc_1^2 + c_2^2 \sum_{i=1}^{k} \int_{-\infty}^{\infty} \left\{ \prod_{\substack{j=1 \\ j\neq i}}^{k} \Phi(x+d+\delta_{i,j}c_1/\delta^*) - \prod_{\substack{j=1 \\ j\neq i}}^{k} \Phi(x-d+\delta_{i,j}c_1/\delta^*) \right\} d\Phi(x) \right\} / k\hat{c}^2 \qquad (5.3)$$

where we employ in (5.3) the $(\hat{c}_1,\hat{c}_2,\hat{d})$-values of the particular procedure P_2 being compared to P_1. The value of $\hat{c}$ which is the solution to (5.2) has been tabulated for selected k and P^* by Bechhofer (1954), Gupta (1963), and Milton (1963); Bechhofer's $\lambda = \hat{c}$, Gupta's and Milton's $H = \hat{c}/\sqrt{2}$.

Table III gives computed RE-values for $RE(P_1:P_2(E))$, $RE(P_1:P_2(C_1))$, and $RE(P_1:P_2(C_3))$ for $k = 2$ and selected P^*, while Table IV gives analogous values for $RE(P_1:P_2(C_1))$ for $k \geq 3$. The computed values given in Tables III and IV were obtained using the $(\hat{c}_1,\hat{c}_2,\hat{d})$-values listed in Tables I and II, respectively. Table V is an abbreviated one which permits comparison of $P_2(C_2)$ and $P_2(C_3)$ with $P_2(C_1)$ via $RE(P_1:P_2(C_i))$ $i = 1,2,3$ for selected extreme (k,P^*)-combinations.

Since P_1 is a special case of $P_2(E)$ and $P_2(C_i)$ $i = 1,2,3$ (see Remark 3.2), it follows that for $k = 2$ we have

$$1 > \max\{RE(P_1:P_2(E)),\ RE(P_1:P_2(C_i)) \quad i = 1,2,3\}, \qquad (5.4)$$

and for $k \geq 3$ we have

$$1 > \max\{RE(P_1:P_2(C_i)) \quad i = 1,2,3\} \qquad (5.5)$$

for all $\underline{\mu} \in \Omega$. Thus, our two-stage procedures $P_2(E)$ and $P_2(C_i)$ $(i = 1,2,3)$ are <u>uniformly</u> (in $\underline{\mu}$) better than the corresponding single-stage procedure P_1 when all guarantee the same probability requirement (2.1). Moreover, when the constraint in (2.2) is replaced by $A(\delta^*) \geq P^*$, and then by $B(\delta^*) \geq P^*$, and finally by $C(\delta^*) \geq P^*$, we have as a consequence of (3.3) that the set of feasible values of (c_1,c_2,d) decreases at each step. Thus the corresponding minima of the objective function of (2.2) increase at each step, and we have for $k \geq 3$:

$$RE_{EMC}(P_1:P_2(C_1)) < RE_{EMC}(P_1:P_2(C_2)) < RE_{EMC}(P_1:P_2(C_3)) \qquad (5.6)$$

TABLE III

Efficiency of P_1 Relative to $P_2(E)$, $P_2(C_1)$, and $P_2(C_3)$ for k = 2 when $\mu_{[2]} - \mu_{[1]} = \delta$

	$\delta = 0$ (EMC)			$\delta = 0.5\delta^*$			$\delta = \delta^*$ (LFC)			$\delta = 4\delta^*$			$\delta = \infty$		
	$P_2(E)$	$P_2(C_1)$	$P_2(C_3)$	$P_2(E)$	$P_2(C_1)$	$P_2(C_3)$	$P_2(E)$	$P_2(C_1)$	$P_2(C_3)$	$P_2(E)$	$P_2(C_1)$	$P_2(C_3)$	$P_2(E)$	$P_2(C_1)$	$P_2(C_3)$
0.9999	0.900	0.907	0.919	0.797	0.796	0.795	0.747	0.737	0.728	0.745	0.735	0.726	0.745	0.735	0.726
0.9995	0.894	0.901	0.920	0.800	0.802	0.805	0.735	0.730	0.715	0.729	0.723	0.706	0.729	0.723	0.706
0.999	0.891	0.900	0.922	0.801	0.807	0.811	0.728	0.726	0.709	0.719	0.716	0.694	0.719	0.716	0.694
0.99	0.878	0.895	0.937	0.814	0.827	0.854	0.722	0.725	0.716	0.683	0.677	0.636	0.683	0.677	0.636
0.95	0.865	0.898	0.972	0.827	0.858	0.931	0.746	0.771	0.827	0.641	0.646	0.559	0.641	0.646	0.559
0.90	0.858	0.905	0.992	0.833	0.879	0.978	0.772	0.816	0.933	0.620	0.644	0.536	0.620	0.644	0.521
0.85	0.853	0.915		0.836	0.898		0.791	0.856		0.606	0.665		0.604	0.662	
0.80	0.850	0.922		0.838	0.911		0.806	0.881		0.605	0.670		0.592	0.654	
0.75	0.847	0.930		0.840	0.923		0.818	0.905		0.622	0.734		0.582	0.701	
0.70	0.845	0.938		0.841	0.934		0.827	0.923		0.662	0.790		0.575	0.720	
0.65	0.844	0.949		0.841	0.947		0.833	0.941		0.720	0.848		0.572	0.689	
0.60	0.842	0.958		0.841	0.957		0.838	0.955		0.779	0.913		0.567	0.700	
0.55	0.842	0.980		0.842	0.979		0.841	0.979		0.825	0.970		0.566	0.767	

where EMC refers to the configuration $\mu_{[1]} = \mu_{[k]}$. This implies that not only is $P_2(C_1)$ U-minimax among our two-stage procedures based on the conservative lower bound $A(\delta^*) \geq P^*$, but also that it is U-minimax among our two-stage procedures based on the conservative lower bounds $B(\delta^*) \geq P^*$ or $C(\delta^*) \geq P^*$ as well. These findings and ones described in the sections below demonstrate that $P_2(C_1)$ is highly effective as a selection procedure with screening.

5.1 $P_2(E)$, $P_2(C_1)$ and $P_2(C_3)$ vs. P_1 for $k = 2$

We have discussed $RE(P_1:P_2(E))$ and $RE(P_1:P_2(C_3))$ in Section 9.1 of T-B. From Table III we note that the performance of $P_2(C_1)$ is always "intermediate" to that of $P_2(E)$ and $P_2(C_3)$ at $\delta = 0$. The range of $\delta > 0$ values for which $P_2(C_1)$ is intermediate depends on P^*; thus, e.g., our computations indicate that for $\delta = \infty$ we have $P_2(C_1)$ as intermediate when $P^* \geq 0.99$ but as poorest when $P^* \leq 0.95$.

5.2 $P_2(C_1)$ vs. P_1 for $k \geq 3$

The performance of $P_2(C_1)$ relative to that of P_1 can be studied using the RE-values given in Table IV. We note that for fixed k and P^*, RE is a decreasing function of the differences $\mu_{[i]} - \mu_{[i-1]}$ $(1 \leq i \leq k-1)$; thus $P_2(C_1)$ capitalizes on favorable configurations of the μ_i $(1 \leq i \leq k)$. The columns headed $\mu_{[k]} = \mu_{[1]}$ and $\mu_{[k]} - \mu_{[k-1]} = \infty$ represent the minimum and maximum RE (measured in terms of $E_{\underline{\mu}}\{T|P\}$), respectively, achieved when $P_2(C_1)$ is used in place of P_1; as noted earlier, based on this criterion the experimenter always gains using $P_2(C_1)$.

We also note that for fixed $\underline{\mu}$ and P^*, RE is a decreasing function of k. Thus the screening feature of $P_2(C_1)$ becomes more effective as k increases.

In the range of P^*-values for which computations were made (these being the ones of greatest practical interest), we note

TABLE IV

Efficiency of P_1 Relative to $P_2(C_1)$ for $k \geq 3$ when the $\mu_{[i]}$ $(1 \leq i \leq k)$ are in Various Configurations

k	P*		$\mu_{[k]}-\mu_{[k-1]}=\delta^*$ $\mu_{[i]}-\mu_{[i-1]}=\delta$ $(2\leq i\leq k-1)$			
		(EMC)	(LFC)			
		$\mu_{[k]}=\mu_{[1]}$	$\delta/\delta^*=0$	$\delta/\delta^*=1.0$	$\delta/\delta^*=4.0$	$\mu_{[k]}-\mu_{[k-1]}=\infty$
3	0.99	0.863	0.663	0.633	0.633	0.595
	0.95	0.882	0.736	0.665	0.658	0.544
	0.90	0.904	0.818	0.735	0.700	0.500
	0.75	0.921	0.917	0.906	0.798	0.485
4	0.99	0.860	0.668	0.633	0.633	0.610
	0.95	0.876	0.718	0.620	0.617	0.538
	0.90	0.902	0.786	0.654	0.641	0.516
	0.75	0.918	0.907	0.834	0.694	0.476
5	0.99	0.858	0.662	0.622	0.622	0.603
	0.95	0.863	0.703	0.601	0.600	0.543
	0.90	0.873	0.750	0.604	0.596	0.504
	0.75	0.909	0.879	0.710	0.631	0.466
10	0.99	0.807	0.626	0.576	0.576	0.566
	0.95	0.798	0.652	0.542	0.541	0.514
	0.90	0.796	0.677	0.524	0.522	0.480
	0.75	0.808	0.740	0.521	0.510	0.439
15	0.99	0.778	0.609	0.557	0.557	0.550
	0.95	0.756	0.623	0.513	0.513	0.494
	0.90	0.756	0.646	0.498	0.497	0.469
	0.75	0.761	0.693	0.471	0.464	0.414
	0.60	0.805	0.767	0.472	0.450	0.386
25	0.99	0.735	0.584	0.530	0.530	0.526
	0.95	0.712	0.593	0.485	0.485	0.473
	0.90	0.702	0.607	0.466	0.466	0.448
	0.75	0.689	0.631	0.432	0.430	0.399
	0.60	0.708	0.671	0.429	0.421	0.374
	0.45	0.713	0.691	0.401	0.380	0.318

that for fixed $\underline{\mu}$ and k, RE decreases and then increases as P* increases; in fact, RE $\to$ 1 as P* $\to$ 1/k or 1.

5.3 $P_2(C_1)$, $P_2(C_2)$ and $P_2(C_3)$ vs. P_1 for $k \geq 3$

Table V which gives $RE(P_1:P_2(C_i))$ i = 1,2,3 for four "extreme" (k,P*)-combinations, namely, k = 3 and 25 and P* = 0.75 and 0.99, can be used to compare the performances of $P_2(C_i)$ i = 1,2,3 over a considerable portion of the range of practical interest of these parameters. The values for $RE(P_1:P_2(C_1))$ are taken from Table IV of the present paper, those for $RE(P_1:P_2(C_3))$ from Table IV of T-B, and those for $RE(P_1:P_2(C_2))$ were computed just for inclusion in this table.

We first note that for fixed k and P* we have, over the range of (k,P*)-values considered, that

$$RE_{\underline{\mu}}(P_1:P_2(C_1)) < RE_{\underline{\mu}}(P_1:P_2(C_2)) < RE_{\underline{\mu}}(P_1:P_2(C_3)) \qquad (5.7)$$

for all $\underline{\mu} \in \Omega$ (including $\{\underline{\mu} | \mu_{[k]} - \mu_{[k-1]} = \infty\}$). This latter is in contrast to the results for k = 2 where it was found that $RE_{\underline{\mu}}(P_1:P_2(C_1)) > RE_{\underline{\mu}}(P_1:P_2(C_3))$ for $\mu_{[2]} - \mu_{[1]} = \infty$.

For fixed $\underline{\mu}$ and k, and P* $\to$ 1, the RE-values are close for the $P_2(C_i)$ i = 1,2,3; this is so since $\hat{n}_1 \to \infty$ and hence the $P_2(C_i) \to P_1$; however, the effect on the RE-values of P* $\to$ 1 depends critically on k for each of the $P_2(C_i)$. For fixed P* "close to unity" we see that the RE-values of the $P_2(C_i)$ are closer for large k than for small k.

For fixed $\underline{\mu}$ and k, and P* $\to$ 1/k, our computational results indicate that the minimax solution $(\hat{n}_1,\hat{n}_2,\hat{h})$ of (3.9) is such that $\hat{h} \to \infty$ and hence $P_2(C_i) \to P_1$ (i = 1,2,3); thus the three procedures perform similarly. However, the value of P* (call it $\overline{P}_i^*$) at which h becomes large enough so that $P_2(C_i)$ becomes "almost equivalent to" P_1 is such that $\overline{P}_1^* < \overline{P}_2^* < \overline{P}_3^*$. Thus for moderate values of P* we find that $P_2(C_1)$ is superior (in terms of RE) to $P_2(C_2)$ which in turn is superior to $P_2(C_3)$.

TABLE V

Efficiency of P_1 relative to $P_2(C_1)$, $P_2(C_2)$ and $P_2(C_3)$ for k = 3 and 25 and P* = 0.99 and 0.75 when the $\mu_{[i]}$ $(1 \leq i \leq k)$ are in various configurations

k	P*	(EMC) $\mu_{[k]} = \mu_{[1]}$			$\mu_{[k]}-\mu_{[k-1]}=\delta^*$, $\mu_{[i]}-\mu_{[i-1]}=\delta$ $(2\leq i\leq k-1)$						$\mu_{[k]} - \mu_{[k-1]} = \infty$		
					$\delta/\delta^* = 0$ (LFC)			$\delta/\delta^* = 1.0$					
		$P_2(C_1)$	$P_2(C_2)$	$P_2(C_3)$	$P_2(C_1)$	$P_2(C_2)$	$P_2(C_3)$	$P_2(C_1)$	$P_2(C_2)$	$P_2(C_3)$	$P_2(C_1)$	$P_2(C_2)$	$P_2(C_3)$
3	0.99	0.863	0.932	0.933	0.663	0.718	0.718	0.633	0.691	0.691	0.595	0.657	0.657
3	0.75	0.921	0.999	1.000-	0.917	0.991	0.995	0.906	0.972	0.981	0.485	0.518	0.535
25	0.99	0.735	0.747	0.747	0.584	0.591	0.591	0.530	0.537	0.538	0.526	0.533	0.533
25	0.75	0.689	0.739	0.768	0.631	0.672	0.695	0.432	0.456	0.479	0.399	0.421	0.445

6. NUMERICAL EXAMPLE

In this section we give a numerical example to illustrate the use of Tables II and IV for $P_2(C_1)$. In addition, we compare the performance (in terms of $E_{\underline{\mu}}\{T|P\}$) of $P_2(C_1)$ with that of the single-stage procedure P_1 and the open sequential procedure $P_S(BKS)$ of Bechhofer, Kiefer and Sobel (1968, Section 12.6.1.1) which samples a vector-at-a-time. The example will show in a striking way the trade-offs that the experimenter has at his disposal when choosing among procedures that guarantee (2.1).

Suppose that $k = 10$, $\sigma = 10$ and that the experimenter specifies $\delta^* = 2$, $P^* = 0.90$; we then anticipate large sample sizes since for $k = 10$ the specification $\delta^*/\sigma = 0.2$, $P^* = 0.90$ is a demanding one.

a) To determine the constants necessary to implement $P_2(C_1)$, we obtain from Table II: $\hat{c}_1 = 2.067$, $\hat{c}_2 = 2.507$, $\hat{d} = 1.342$. Using (3.11) these yield $\hat{n}_1 = [(2.067/0.2)^2] = [106.8] = 107$, $\hat{n}_2 = [(2.507/0.2)^2] = [157.1] = 158$, $\hat{h} = 1.342(2)/2.067 = 1.299$.

b) To determine the constant necessary to implement P_1, we obtain from Table I of Bechhofer (1954) that $\hat{c}$ of our (5.2) is 2.9829. Thus $n = [(2.9829(10)/2)^2] = [222.4] = 223$, which is the number of observations required from <u>each</u> of the 10 populations.

c) We obtain estimates of $E_{\underline{\mu}}\{T|P_S(BKS)\}$ from Tables 18.4.5 and 18.4.10 of BKS (1968) for the LFC and the EMC, respectively; these are 1453 and 2906, respectively.

The above results are summarized in Table VI.

The entries in Table VI illustrate the savings in $E_{\underline{\mu}}\{T|P\}$ when $P_2(C_1)$ is used in place of P_1. In addition, the entries for $P_S(BKS)$ show the dramatic further savings that can be achieved using that procedure if sequential sampling is a viable method of experimentation for the practical problem at hand <u>and</u> it is anticipated that the largest population means are not too close. However, it must be emphasized here that we are presently

TABLE VI

$E_{\underline{\mu}}\{T|P\}$ for P_1, $P_2(C_1)$, and $P_S(BKS)$
for Selected Configurations of the Population Means
when $k = 10$, $\delta^*/\sigma = 0.2$, $P^* = 0.90$

Procedure	$E_{\underline{\mu}}\{T\|P\}$		
	EMC	LFC	$\mu_{[10]} - \mu_{[9]} = \infty$
P_1	2230	2230	2230
$P_2(C_1)$	1775[a/]	1510[a/]	1070[a/]
$P_S(BKS)$	2906	1453	10

Note: The entries for $P_2(C_1)$ in Table VI were computed by multiplying the relevant relative efficiencies in Table IV by 2230.

focusing on $E\{T\}$. The distribution of T has a very large standard deviation for $P_S(BKS)$ when $E\{T\}$ is large (see Tables 18.4.5 and 18.4.10 of BKS (1968) where for $P^* = 0.90$ the estimated standard deviations of T are given as 713.4 and 1744.6 for the LFC and EMC, respectively) and is highly skewed to the right; thus if the experimenter uses this procedure he must be prepared to accept occasional very large values of T. The closed sequential procedure of Paulson (1964) (with the improvement of Fabian (1974)) which samples a vector-at-a-time and eliminates populations is superior to $P_S(BKS)$ in terms of $E_{\underline{\mu}}\{T|P\}$ over certain ranges of values of k and P^*; for the problem at hand with $k = 10$, $P^* = 0.90$, $\delta^*/\sigma = 0.2$ and Paulson's design parameter $\lambda = \delta^*/2$ we have, using Fabian's improvement, that the maximum number of stages to terminate experimentation for Paulson's procedure is 381, and hence an upper bound on $E_{\underline{\mu}}\{T|P\}$ which is extremely conservative

(since it does not take into account the fact that populations are permanently eliminated prior to termination of experimentation) is 3810. Here again the problem at hand must be such that a completely sequential procedure is feasible, and the experimenter must be prepared to accept very large values of T.

7. CONCLUDING REMARKS

We feel that in spite of the relatively heavy financial costs involved in obtaining the design constants necessary to implement $P_2(C_1)$, they were justified by the final results. For we have been able to demonstrate conclusively that $P_2(C_1)$ represents a significant improvement over both $P_2(C_2)$ and $P_2(C_3)$ (as well as over P_1). And we now can offer a highly effective two-stage procedure, incorporating screening, which is easy to implement.

8. DIRECTIONS OF FUTURE RESEARCH

In Section 10 of T-B we postulated several unsolved problems associated with P_2. All of these still remain open problems. The most important of these (at least from a theoretical point of view) is that of determining the LFC of the μ_i for $k > 2$. If the conjectured LFC $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ can be proved to be the true one, and if an efficient algorithm can be found for evaluating the exact $P_{\underline{\mu}}\{CS|P_2\}$ for $\underline{\mu}$ in the LFC, then it would be of considerable interest to determine how much decrease in $\sup_{\mu\in\Omega} E_{\underline{\mu}}\{T|P\}$ can be achieved if $P_2(E)$ is used in place of $P_2(C_1)$ for $k \geq 3$.

9. ACKNOWLEDGMENTS

The authors would like to thank Dr. Roy Milton and Professors Shanti Gupta and Donald Owen for helpful suggestions concerning some problems encountered in computing the tables given in this paper.

This research was supported by NSF Grant ENG 77-06112 at Northwestern University, and U.S. Army Research Office--Durham contract DAAG29-77-C-0003 and Office of Naval Research contract N00014-75-C-0586, both at Cornell University.

BIBLIOGRAPHY

Abadie, J. and Carpentier, J. (1969). Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints. Optimization (ed. R. Fletcher), New York, Academic Press.

Abramowitz, M. and Stegun, I.A. (1964). Handbook of Mathematical Functions. Nat. Bureau of Standards, Appl. Math. Ser. 55, Washington, D.C.: U.S. Govt. Printing Office.

Alam, K. (1970). A two-sample procedure for selecting the population with the largest mean from k normal populations. Ann. Inst. Statist. Math. 22, 127-36.

Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 25, 16-39.

Bechhofer, R.E., Kiefer, J. and Sobel, M. (1968). Sequential Identification and Ranking Procedures. Chicago, Ill.: The University of Chicago Press.

Bechhofer, R.E. and Tamhane, A.C. (1974). An iterated integral representation for a multivariate normal integral having block covariance structure. Biometrika 61, 615-19.

Borth, D.M. (1973). A modification of Owen's method for computing the bivariate normal integral. J. R. Statist. Soc. C, 22, 82-85.

Cadwell, J.H. (1951). The bivariate normal integral. Biometrika 38, 475-79.

Fabian, V. (1974). Note on Anderson's sequential procedures with triangular boundary. Ann. Statist. 2, 170-6.

Gupta, S.S. (1963). Probability integrals for multivariate normal and multivariate t. Ann. Math. Statist. 34, 792-828.

Milton, R.C. (1963). Tables of the equally correlated multivariate normal probability integral. Tech. Rep. No. 27, Dept. of Statistics, Univ. of Minnesota, Minneapolis, Minnesota.

Owen, D.B. (1956). Tables for computing the bivariate normal probabilities. Ann. Math. Statist. 27, 1075-90.

Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. Ann. Math. Statist. 35, 174-80.

Somerville, P.N. (1974). On allocation of resources in a two-stage selection procedure. Sankhya B, 36, 194-203.

Tamhane, A.C. (1975). On minimax multistage elimination type rules for selecting the largest normal mean. Tech. Report 259, Dept. of Operations Res. Cornell Univ., Ithaca, N.Y.

Tamhane, A.C. and Bechhofer, R.E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. Commun. Statist. A6, 1003-33.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #377	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A TWO-STAGE MINIMAX PROCEDURE WITH SCREENING FOR SELECTING THE LARGEST NORMAL MEAN (II): AN IMPROVED PCS LOWER BOUND AND ASSOCIATED TABLES		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ajit C. Tamhane and Robert E. Bechhofer		8. CONTRACT OR GRANT NUMBER(s) NSF ENG 77-06112 DAAG29-77-C-0003 N00014-75-C-0586
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Operations Research & Industrial Engineering, College of Engineering Cornell University, Ithaca, N.Y. 14853		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS National Science Foundation Washington, D.C. 20550		12. REPORT DATE June 1978 (Revised Oct. 1978) 13. NUMBER OF PAGES 26
14. CONTROLLING OFFICE NAME AND ADDRESS Sponsoring Military Activities: U.S. Army Research Office, P.O. Box 12211, Research Triangle Park, N.C. 27709 and Statistics and Probability Program, Office of Naval Research, Arlington, VA 22217		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Selection procedure; ranking procedure; minimax procedure; two-stage procedure; indifference-zone approach; screening

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper is a follow-up to an earlier article by the authors in which they proposed a two-stage procedure with screening to select the normal population with the largest population mean when the populations have a common known variance. The two-stage procedure has the highly desirable property that the expected total number of observations required by the procedure is always less than the total number of observations required by the corresponding single-stage procedure of Bechhofer (1954), regardless of the configuration of the population means. The present paper contains new results which make possible the more

2

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

efficient implementation of the two-stage procedure. Tables for this purpose are given, and the improvements achieved (which are substantial) are assessed.